Herpetological Review, 2006, 37(1), 57–60.

Addressing Error in Identification of *Ambystoma maculatum* (Spotted Salamanders) Using Spot Patterns

EVAN H. CAMPBELL GRANT*
and
PRIYA NANJAPPA
USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road
Laurel, Maryland 20708, USA

**Corresponding author; e-mail: ehgrant@usgs.gov*

Numerous capture-recapture studies of amphibians have used the unique color patterns of certain amphibian species to "mark" individuals, recording pattern information with sketches, photographs, or a coding system (Andreone 1986; Forester 1977; Gill 1978; Kurashina et al. 2003; Marvin 1996; Nace et al. 1973; Nijhuis and Kaplan 1998; Tilley 1980). Such non-invasive techniques are preferred over physical alteration of an animal, which may include toe-clipping, visual implant elastomer, or other added marks (Bailey 2004; McCarthy and Parris 2004; Murray and Fuller 2000; but see Funk et al. 2005). Regardless, bias associated with any individual identification method should be assessed and reported, though this has been infrequent in amphibian capture-recapture studies (but see Bailey 2004; McCarthy and Parris 2004; Muths et al. 2000).

For long term capture-recapture studies to be successful, marks must not be lost over time; this is assumed to be true for natural patterns that serve as marks (Gill 1978). More importantly, marks should not be overlooked or misidentified by observers. Reduction of observer bias in mark recognition is paramount in studies that estimate population sizes and document long term trends.

The goals of our study were to: 1) quantify error rates in individual identification through specific pattern mapping approaches; 2) address sources of error by developing a modified pattern mapping method; 3) compare bias associated with these two methods and identify which method minimizes observer bias in both pattern mapping in the field and individual identification in the lab; and 4) determine the search effort necessary to find all individuals in a dataset. We assessed bias in pattern matching of the Spotted Salamander (*Ambystoma maculatum*), a mole salamander with conspicuous and distinctive yellow spot patterns.

Materials and Methods.—We used data from March–April 2002 and 2004 for this study. To capture migrating spotted salamander adults we used artificial cover objects (61 × 122 cm plywood boards) and a drift fence completely encircling Laura's Pond, a semipermanent fishless pond in Beltsville, Maryland (USA). We digitally photographed the dorsal patterns of each salamander on a flat surface using a Nikon® Coolpix 995 digital camera, in the shade using the integrated flash to reduce glare. We did not sedate or restrain salamanders.

Between 03 March and 08 April 2002, we employed a multi-parameter (MP) method of pattern mapping (Table 1) based on a modification of the Loafman method (Loafman 1991). All pattern mapping was done in the field. We counted any spot > 1 mm spanning two body areas once (in the area which contained the majority of the spot). Twenty observers (2–10 per visit) participated in recording pattern data for 654 salamander captures. We entered and sorted all pattern data in Microsoft® Excel™ spreadsheets; each record represented one capture occasion and was associated with its digital photograph. We used Microsoft® Photo Editor™ to view multiple photographs at once. In the lab, we sorted spot counts by the six body areas (Table 1). Individuals were identified by comparing and matching digital photographs among these grouped records. Of 378 individuals identified, 136 were captured more than once. We noted that for some of these individuals, recorded pattern data varied among capture occasions. Because we had digital photos from each capture, we were able to verify that discrepancies were due not to pattern changes, but rather to errors in field data collection using the MP method.

We tested the hypothesis that the variation in field data recorded using the MP method differed among areas of the body (e.g., head, each leg, and body + tail). We selected at random a set of 50 individuals with multiple captures (mean number of captures per individual = 3.82, range = 2–13), and assessed the variation in recorded spot counts by body area between observations of each individual. We hypothesized that head pattern data would have the lowest observer bias based on Doody (1995), who observed that head patterns often were sufficient to distinguish individual Spotted Salamanders.

We compared the efficiency of the MP method to a reduced parameter (RP) method of pattern mapping (Table 1) which we hypothesized would facilitate pattern mapping in the field and individual identification in the lab. As in the MP method, a spot spanning two body areas was counted only once. Other supporting pattern data also were recorded, including a categorical assessment of the spots anterior to the gular fold [i.e., eye code, where 0 = no spots by either eye, 1 = spot(s) by one eye only, 2 = spot(s) by both eyes; see Nace et al. 1973], eye spot symmetry (yes/no) and gular spot symmetry (yes/no), where symmetry describes the same number of spots in mirror orientation across an anterior–posterior midline.

To assess laboratory error, we tested whether the method of pattern mapping (MP vs. RP) influenced the ability of observers to identify individuals by matching pattern data and associated photographs. We also tested whether the ability of observers to identify matches varied based on the number of records. Two sets each of 50 and 100 capture records were chosen at random from the 2002 (MP) data; individuals were manifested in these "test datasets" as one record (one capture), or as a set of >1 records (multiple captures or "matches"). The number of matches per dataset was known; individuals were identified and verified in the full set of pattern data. Based on the associated photos, a copy of each file was then altered by one of the authors (EHCG) to show only the parameters pertaining to the RP method (Table 1). Five observers were presented with the same 8 test datasets (2 MP and 2 RP datasets each of 50 and 100 records).

The observers identified matching records in the lab, recording elapsed time between finding matches, as well as total time spent on each test dataset. We created an "observer performance" matrix (modeled after a traditional capture history matrix) where rows represented each salamander capture record, and columns represented the classification results from each observer for each salamander. Classifications were designated as either "correct" (matching records or unique records correctly identified as such) and received a value of "1," or "incorrect" (failure to detect matching records when present) and received a value of "0." We combined the observer matrices from all test datasets for each method, and tested whether the binomial probability of correct classification differed between: 1) the two methods (MP vs. RP); 2) the number of records presented to each observer (50 or 100); and 3) observers, using the program CONTRAST (Hines and Sauer 1989). We also calculated the average time to find a match for each method, and estimated the number of searches through a full dataset necessary to identify all matches correctly by $(1-p)^t$, where p is the probability of correct classification and t is the number of searches by one observer.

TABLE 1. Descriptions of two pattern mapping methods used to distinguish individual Spotted Salamanders, *Ambystoma maculatum*.

	Method	
Character	Multi-parameter (MP)	Reduced parameter (RP)
Spot definition	Any yellow or orange spot >1.0 mm maximum length or width	Any yellow or orange spot, regardless of size
Body areas counted	Six body areas: 1) total head spot count, 2) total body (torso+tail) spots 3–6) separate counts for each leg	Two body areas: 1) total head spot count, 2) sum of counts from the front legs
Description of head spot pattern	No additional descriptors for head spot location	Additional descriptor codes for head spot location and symmetry

We employed the RP method during 08 March to 14 April 2004 because fewer parameters were collected, thereby reducing field handling time per salamander. We recorded a total of 592 capture occasions of Spotted Salamanders. One observer searched the data for matches, eliminating all secondary captures from the data set (leaving only presumed initial captures and those records representing salamanders presumed to have been captured only once) before searching again. We counted the actual number of search iterations necessary to identify all individuals and compared this to our estimated number of iterations based on the average classification probability of five observers from the test datasets.

Results.—Within the 50 Spotted Salamander individuals with multiple captures from the preliminary assessment of the MP method, the mean and variation in recorded spot counts for the head (mean = 0.48, var = 0.51, range = 0–5 spots) and legs (mean = 0.88, var = 0.58, range = 0–3 spots) were smaller across capture occasions than for the body (mean = 1.48, var = 1.01, range = 0–6 spots).

Comparing the MP and RP methods in the test datasets, the probability of correct classifications was high for all observers between both method [p_{MP} = 0.97 (SE = 0.005), p_{RP} = 0.96 (SE = 0.005); χ^2 = 0.88, P = 0.35] and the number of records presented to the observers [p_{50} = 0.97 (SE = 0.006), p_{100} = 0.96 (SE = 0.004); χ^2 = 0.51, P = 0.48]. Observers did not differ in their ability to classify records correctly (mean $p_{test\ datasets}$ = 0.96; χ^2 = 5.86, P = 0.21), although actual correct classification varied (MP: mean = 80 + 23 %, range = 33 to 100 %; RP: mean = 74 + 27 %, range = 0 to 100 %). Of 1500 total records (5 observers × 2 sets × [100 + 50 salamanders]) in test datasets per method, observers incorrectly classified 52 (3.5%) MP and 60 (4.0%) RP records. The amount of time spent searching for matches did not explain the proportion of correct classifications (Fig. 1). Assuming that the mean observer detection probability (p) does not change as observers find more matches in a dataset, we estimated that a minimum of four searches through the data are necessary to identify all individuals (Fig. 2).

Comparing these results to our 2004 data collected using the RP method, we found that the actual probability of correct classification was less than our estimate (p_{2004} = 0.90 vs. $p_{test\ datasets}$ = 0.96, Fig. 2), and the number of iterative searches necessary was greater than our estimate (t = 8 searches, Fig. 3). The probability of incorrect classification decreased over the 8 searches (1-p; Fig. 3).

Discussion.— The MP and RP methods were comparable in efficiency and accuracy in identifying individuals in the lab, and the probability of correct classification was independent of method, number of records, or observer. This suggests that the method of pattern mapping does not affect the ability of observers to match individuals in small datasets, provided that supporting data allows records to be grouped. Observer performance, measured by number of incorrect classifications, was slightly better using the MP

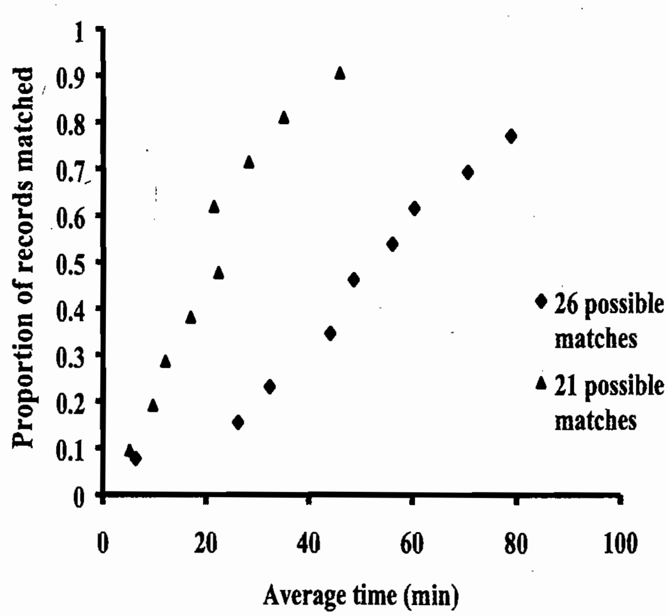

FIG. 1. The proportion of records matched (where a "match" is defined as >1 records representing multiple captures of an individual) vs. time spent by 5 independent observers using Spotted Salamander (*Ambystoma maculatum*) pattern data from Laura's Pond in Beltsville, Maryland, USA. Each series represents one of two replicates from the 2002 test datasets (MP method) with 100 individuals each. One dataset had 21 possible matches, while the other had 26 possible matches. Not surprisingly, a greater number of potential matches present in a data set requires a longer amount of time to identify the complete set of matches.

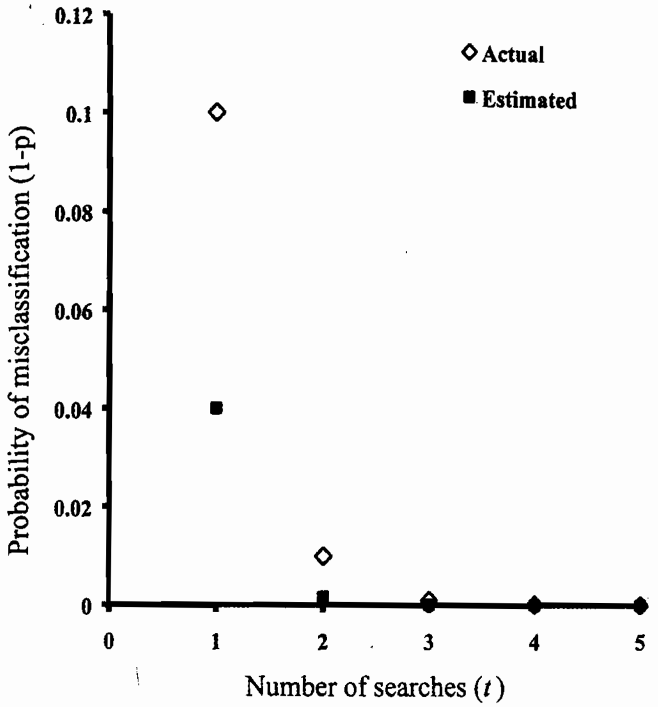

FIG. 2. The estimated (using test datasets) and actual probability (using 2004 field data) of incorrect classification (1-p), or missed matches, in Spotted Salamander pattern data. The number of searches represents iterations necessary for one observer to search a data set and identify all possible individuals, estimated by $(1\text{-p})^t$, where p is the probability of classifying a record correctly and does not change, and t is the number of searches. The mean value $p_{\text{test datasets}} = 0.96$ was estimated from 5 independent observers searching 8 test datasets, and the value $p_{2004} = 0.90$ represents the actual classification probability from the first iterative search of the 2004 RP data collected in the field.

method. However, we observed that the RP method reduced the time to handle and record pattern data in the field. This likely reduced stress on the animal, particularly those captured multiple times, and therefore we prefer the RP method.

Inconsistent pattern mapping in the field (field error) prevented some records representing the same individual from becoming grouped when sorted in the spreadsheet, and thus matches within these groups may have been missed. Variation in pattern data over multiple captures of the same individual may have partially resulted from our size-restricted "spot definition" (Table 1) in the MP method; this restriction was removed in the RP method. Field error due to transpositions of spot counts for right versus left or front versus rear leg were eliminated in the RP method by summing spot counts for both front legs. Out-of-focus pictures, poor lighting/contrast, and photos that did not show the entire pattern clearly also could have resulted in missed matches. Problems in field photography were not addressed here, but could be eliminated by using standardized photographic procedures (e.g., Doody 1995; Ravela and Gamble 2004).

To identify all individuals from capture records, we found that multiple searches through the data by one observer were necessary. Eliminating one record from each matching pair in the data set after each search reduces the number of capture records to compare. The number of individuals represented by multiple records in a data set will naturally affect the amount of time necessary to detect all individuals (Fig. 1). In our 2004 data set, 8 iterations of sorting and matching records were necessary to identify all individuals (Fig. 3). However, the presence of a small amount of error in classification may not severely affect the estimates of population size (Miller et al. 2002). An assessment of the relative contribution of the two types of error investigated here (i.e., field error vs. laboratory error resulting from missed matches in the dataset) can contribute to decisions regarding whether a small amount of error in detecting matching records in the lab can be tolerated. Our results suggest that the probability of false matches (classifying records as matches when they are not) is probably small. The probability of incorrect classification fell below p = 0.02 after 3 searches through the data set (Fig. 3), suggesting that fewer searches may not appreciably affect population estimates.

Previous studies have used counted parameters alone (Loafman 1991) or photographs alone (Bailey 2004; Kurashina et al., 2003) to identify individuals using pattern data. Incorporating supporting information with photos or drawings helps to increase the efficiency of identifying individuals (Gill 1978; Nace et al. 1973). We summarize pattern data in spreadsheet software, which allows us to sort the data and group records with similar characteristics (e.g., number of head spots), and then compare photos within groupings. In studies where relatively large numbers of salamander captures are possible, this combination of data types may facilitate the identification of individual salamanders by narrowing the number of records to compare, and digital photos allow the confirmation of matches. This approach is specific to discrete and enumerable pattern features such as spots, though it might be ap-

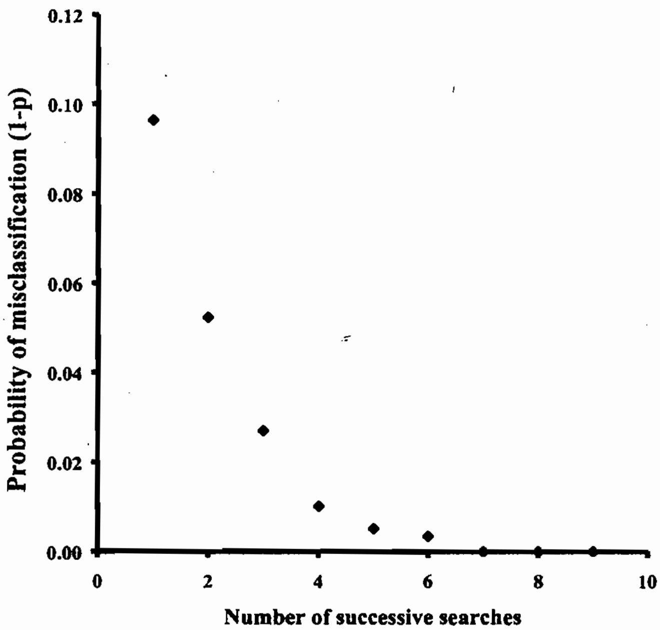

FIG. 3. The observed probability of incorrect classification by the number of searches of the 2004 field-collected data using the RP method. After each search, any records designated as secondary captures were removed from the data set prior to a subsequent search (leaving only presumed initial captures and those salamanders presumed to be observed only once). The probability of incorrect classification (i.e., missed matches) decreases with successive searches through the data. All individuals were identified by the 8th iteration of the successive searches; the 9th search confirmed that no additional matching records were present.

plicable to other elements of pattern. Future advances in technology (i.e., ambystomatid pattern-recognition software, Ravela and Gamble 2004; D. Church, pers. comm.) will automate the process of identifying individuals. Pattern recognition research is progressing in the field of facial recognition (Zhao et al. 2000), and this emerging body of literature (and associated products) can have practical applications for wildlife biology. Even with automated pattern recognition, a subset of computer-determined individuals may require validation using methods similar to the process we present (see also Whitehead 1990; whale fluke identification). Regardless of method, some assessment of bias is useful in evaluating whether a technique violates the assumptions of capture-recapture modeling (specifically 1) that marks are not lost during the period of study and 2) observers can recognize marked individuals, and do not designate marked individuals as new captures). Validation of data subsets can allow investigators to evaluate a marking approach, and thereby qualify the derived estimates of population size or trend.

Acknowledgments.—This project was supported by the USGS Amphibian Research and Monitoring Initiative, Northeast region. Brand names provided in this paper are for information only and do not constitute endorsement by the US Geological Survey. Our manuscript was greatly improved by comments from L. Bailey, S. Doody, L. Gamble, M. Mazerolle, D. Olson, C. Phillips, and two anonymous reviewers. Many thanks to the observers including R. E. Jung, A. Mongeon, T. Plenderleith, and E. Wilson. We also thank W. Crouch for insightful discussions, L. Bailey for suggestions on study design and data analysis, and the late Laura Mazanti for initiating the study of Spotted Salamanders at this study site.

Literature Cited

Andreone, F. 1986. Considerations on marking methods in newts, with particular reference to a variation of the "belly pattern" marking technique. Bull. Brit. Herpetol. Soc. 16:36–37.

Bailey, L. L. 2004. Evaluating elastomer marking and photo identification methods for terrestrial salamanders: marking effects and observer bias. Herpetol. Rev. 35:38–41.

Doody, J. S. 1995. A photographic mark-recapture method for patterned amphibians. Herpetol. Rev. 26:19–21.

Forester, D. C. 1977. Comments on the female reproductive cycle and philopatry by *Desmognathus ochrophaeus* (Amphibia, Urodela, Plethodontidae). J. Herpetol. 11:311–316.

Funk, W. C., M. A. Donnelly, and K. R. Lips. 2005. Alternative views of amphibian toe-clipping. Nature 433:193.

Gill, D. E. 1978. The metapopulation ecology of the red-spotted newt, *Notophthalmus viridescens* (Rafinesque). Ecol. Monogr. 48:145–166.

Hines, J. E., and J. R. Sauer. 1989. Program CONTRAST: a general program for the analysis of several survival or recovery rate estimates. Fish and Wildlife Technical Report 24, Washington, DC.

Kurashina, N., T. Utsunomiya, Y. Utsunomiya, S. Okada, and I. Okochi. 2003. Estimating the population size of an endangered population of *Rana porosa brevipoda* Ito (Amphibia: Ranidae) from photographic identification. Herpetol. Rev. 34:348–349.

Loafman, P. 1991. Identifying individual spotted salamanders by spot pattern. Herpetol. Rev. 22:91–92.

Marvin, G. A. 1996. Life history and population characteristics of the salamander *Plethodon kentucki* with a review of *Plethodon* life histories. Am. Midl. Nat. 136:385–400.

McCarthy, M. A., and K. M. Parris. 2004. Clarifying the effect of toe-clipping on frogs with Bayesian statistics. J. Appl. Ecol. 41:780–786.

Miller, C. R., P. Joyce, and L. P. Waits. 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160:357–366.

Murray, D. L., and M. R. Fuller. 2000. A critical review of the effects of marking on the biology of vertebrates. *In* L. Boitani and T. K. Fuller (eds.), Research Techniques in Animal Ecology: Controversies and Consequences, pp. 15–64. Columbia University Press, New York.

Muths, E., P. S. Corn, and T. R. Stanley. 2000. Use of oxytetracycline in batch-marking post-metamorphic boreal toads. Herpetol. Rev. 31:28–32.

Nace, G. W., C. M. Richards, and G. M. Hazen. 1973. Information control in the amphibian facility: The use of *Rana pipiens* disruptive patterning for individual identification and genetic studies. Am. Zool. 13:115–137.

Nijhuis, M. J., and R. H. Kaplan. 1998. Movement patterns and life history characteristics in a population of the cascade torrent salamander (*Rhyacotriton cascadae*) in the Columbia River Gorge, Oregon. J. Herpetol. 32:301–304.

Ravela, S., and L. R. Gamble. 2004. On recognizing individual salamanders. Proc. 6th Asian Conference on Computer Vision 2:741–747.

Tilley, S. G. 1980. Life histories and comparative demography of two salamander populations. Copeia 1980:806–821.

Whitehead, H. 1990. Computer assisted individual identification of sperm whale flukes. Report of the International Whaling Commission, Special Issue 12:71–77.

Zhao, W., R. Chellappa, A. Rosenfeld, and P.J. Phillips. 2000. Face recognition: a literature survey. ACM Computing Surveys 35:399–458.